

Title	It's The Books Against The Computers! Who wins?
Author(s)	Nelson, R. William
Citation	大阪外国語大学学報. 70(1) p.41-p.53
Issue Date	1985-11-30
oaire:version	VoR
URL	https://hdl.handle.net/11094/81061
rights	
Note	

Osaka University Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

Osaka University

It's The Books Against The Computers! Who Wins?

William R. NELSON

English Department, Osaka
University of Foreign Studies

Introduction

It wasn't so many years ago that institutions, enamored with the then 'recent' breakthroughs in technology, began fortifying their language departments with all manner of recording devices; never mind the cost...or the fact that so few knew how to use them effectively. The language laboratory had arrived. No such devil-may-care attitude has greeted current breakthroughs, but computers are now being carefully examined as educational devices. A number of firms and individuals have developed educational software; however, I have found little in the literature comparing the effectiveness of computer-aided learning with that of 'good old' book learning.

To help fill that void (although rather modestly) a rote learning task, involving opposing pairs of American English idioms, adjectives, nouns and verbs (see Appendix A for complete inventory) was devised. The task was to learn one hundred pairs, although a number of the 'pairs' consisted of more than two members. Those belonging to the experimental group chose to employ either the book or the computer. Each person worked independently and was tested at predetermined intervals.

For each individual, learning and retention curves were plotted (contact author for original data) and analyzed. Their form indicated that while learning from a book was faster, utilizing a computer resulted in greater retention.

Equipment and Software

An NEC PC-8801 computer with an Epson TF-20 double disk drive was used for random number generation, developing and running the learning program, and, with a Silver-Reed EX 44 unidirectional daisy-wheel typewriter/printer, for word-processing. The PC-8801 is an 8-bit computer with 64K of random access memory (RAM). In 1983 the entire system (including an NEC PC-8853K 14-inch color display) cost ¥600,000 (approx. \$2400). Because of rapid advances each year, an identical system would be much less expensive today.

The idea and the contents of the learning program were developed by the author; however, the computer program itself and the introductory graphics owe their existence to the great help of Mr. Masaki Manabe, an assistant in the Department of Architectural Engineering,

College of Engineering, at Oita University in Kyushu, Japan. Further, the computer programs to handle the statistical work owe their existence to the great help of Mr. Mitsuo Yoshida.

In developing the learning program, first, six hundred opposing pairs of idioms, adjectives, nouns and verbs were collected. Because of disk storage limitations, the list was pruned down to three hundred sixty. The three hundred sixty pairs were then divided into three groups, according to the level of difficulty; however, only the most difficult level was used because it constituted a real learning task. Under normal circumstances, a person using the program could freely choose any of the three levels; during the experiment, the program was fixed so that no matter what choice was made, the advanced program would begin. The scoring procedure was also changed; again, under normal circumstances, players would be rewarded points for speed, thus a correct answer within three seconds would be worth more than a correct answer within ten. This, of course, would make it difficult to generate accurate learning curves, so each correct answer was worth five points only, no matter how long it took to answer.

The program also consisted of a help menu and an editing program that could be used to change the contents of the main learning program. The students (participants were all students of Osaka University of Foreign Studies in Minō, Japan) were not told about the editing program, but they were told about the help menu which they could access using the 'Y' key in response to the question, "Do you want help? (Y/N)". The help menu provided the idiom or word being asked about, its meaning and a sample sentence (constructed in a way so as to make the meaning easy to grasp). In addition, the opposing idioms or words were displayed, along with their meanings and a sample sentence.

The main program begins with some brief introductory graphics which show the name of the 'game' and the name of the author. The prompt, "Do you want instructions? (Y/N)", appears then. If the player needs instructions, he pushes the 'Y' key. Because of the changes mentioned above, the instructions were also altered. The version used for the experiment appears below:

GAME INSTRUCTIONS

This is a game to help you practice opposites in English. Opposites are words like 'up' and 'down', or 'big' and 'small'. When you finish reading these instructions and press the RETURN key, you will be asked what class you want to join. Please push '3' for the Advanced Class, and then press RETURN. You should then choose how many questions you want, and press the RETURN key again. The game will then start.

1. As each word or phrase appears on the screen, you should type in your answer and press the RETURN key.
2. If you don't know the answer, just press the RETURN key, and the correct answer or answers will appear.
3. If you want more help, answer 'Y' to the question, "Do you want help? (Y/N)". When you finish reading, press the RETURN key and the game will continue.

READY? THEN LET'S GO! (press RETURN)

After pressing '3' for the Advanced Class, pressing RETURN, choosing the number of questions wanted, and pressing RETURN again, the player will see the following format on the screen:

GUESS THE OPPOSITE WORD OR PHRASE

Advanced Class

Question 1

The phrase is.....RIGHT SIDE OUT

Your Total Score 0

The opposite is.....?

Perfect Total Score 0

If the player makes a mistake, the following is displayed in the lower left quarter of the screen:

Oh! Sorry!

The correct answer: INSIDE OUT

Do you want help? (Y/N)

Utilizing the computer's internal clock, the program was further set up to make use of both the randomization function and the swapping function which meant that any one of the two hundred thirty-seven items (one hundred 'pairs' as explained above) could appear first. The computer's randomization capability is restricted in the sense that it only randomized the order of appearance of the two hundred thirty-seven items. So, while both or all members of a pair could have appeared during any one session, no item could appear more than once. And just for fun, the color of the words on the screen changed randomly with each new question.

A special room with desks was prepared for the students who chose to learn with a book. The contents of the book were alphabetized according to the first letter of the first word of each item, and each item was cross-referenced, but other than that, the contents exactly mirrored the contents of the help menu mentioned above.

Experimental Design

A group of one hundred sixty-three English students (consisting of fifty-eight freshmen, sixty sophomores, thirty-three juniors and twelve seniors) were given a pre-test. The test consisted of thirty items (12.7 percent of the total) selected at random, utilizing a random number chart generated by the computer. As there was no way to administer the test to all participants at the same time, it was done by classes. Each class was instructed to:

1. write down the opposite of each word or phrase that appeared on the pre-test;
2. answer some questions about their exposure to English;
3. make no special effort to try to remember or think about the items on the test; and
4. to not talk about the project with anyone, including the other members of what would become the experimental or control groups.

The purpose of numbers 3 and 4 was to maintain as much control as possible over the

experiment and later, concerning the two experimental groups, to monitor as closely as possible the total learning time for each. These purposes were explained to all.

After the completion of the pre-test, volunteers were asked to either join the book group or the computer group. Twenty-seven students joined the book group and twenty-five joined the computer group. Their group means as well as that of the control group were checked for any significant differences (t -test); there were none, thus indicating that all participants were starting out at the same level.

Those in the book group were instructed to:

1. only use the prepared dictionaries and notebook (ten copies of the dictionary were made so that as many as ten students could study at the same time; the notebook was provided for them to keep careful track of their study time, to write notes in, and for me to enter their test scores in);
2. only study in the room that was prepared for them;
3. keep track of their study time by entering the date, beginning time and ending time of each study session (they were free to choose their own study times);
4. study using any method they preferred;
5. leave the copy of the dictionary and their notebook in my office, and
6. to not talk about any facet of their learning task with anyone.

In addition, those in the book group were given tests (consisting of ninety-nine to one hundred sixteen items each) after every two hours of study. The type of the test was exactly like that of the pre-test in which they were asked to write down the opposite of each item appearing on the test. After each student reached the 100% level (less in some cases because of time and/or scheduling problems), he or she was given another test and asked to provide the meanings of those items appearing on the test. Altogether, six tests were prepared, all of which varied in order and selection, although there was a great deal of overlap because each of the items was taken from a total pool of two hundred thirty-seven. However, in no case did two opposing members of one pair appear in the same test.

Those in the computer group were taught how to use the computer and the learning program. They were also given notebooks in which to record the date, beginning time and ending time of each study period. In addition to that, they were asked to record their scores as they appeared on the display, that is, their own score plus the figure for what would have been a perfect score had they answered all the questions correctly. They, too, were instructed not to discuss any facet of the learning program with anyone, nor to do any related study outside the computer room. Because their learning task involved a kind of continuous test, they were not given any additional pairing tests. However, when a player's score reached 500 over 500 (player's score over perfect score for one hundred questions), he or she was given a meaning test which was the same as that given to the members of the book group. In this group, too, there were some who did not reach the 100% level for the same reasons mentioned above.

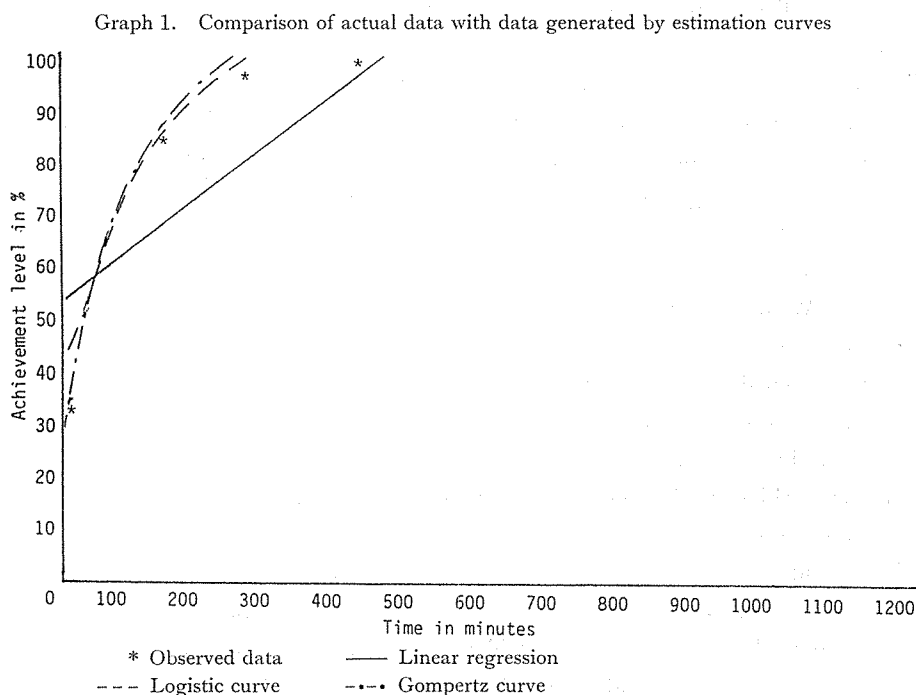
As all the members of both the book group and the computer group completed the learning phase, they were again told not to discuss any facet of the experiment with anyone and not to

try to remember or re-study any of the work they had done. All students were thanked for their effort and time, and told that when I had finished analyzing the data, I would tell them the results (a promise that I will make good on with the completion of this report).

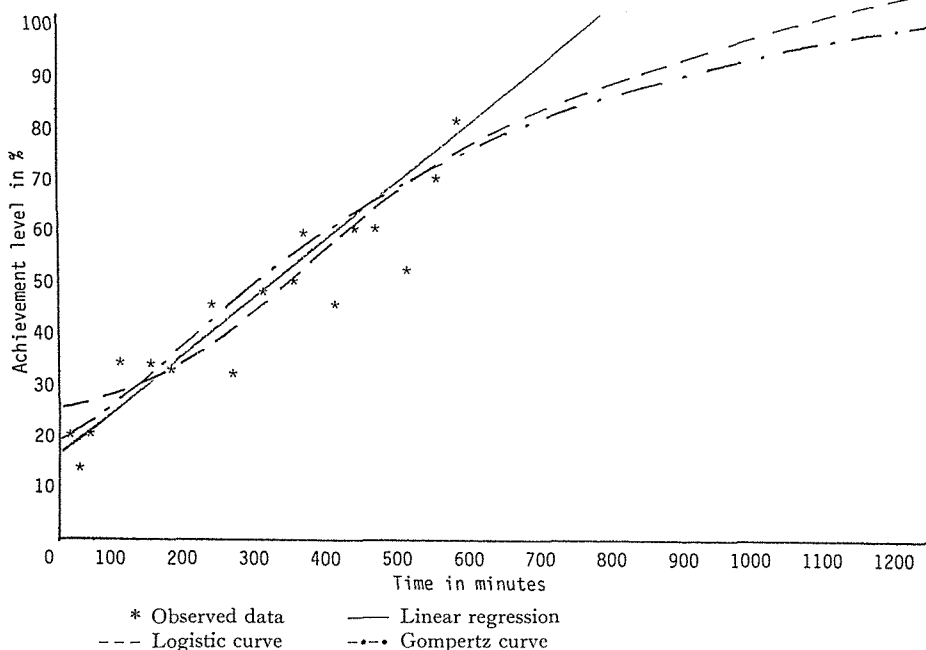
Two more randomly ordered (as explained above) pairing tests of one hundred items each were generated, and these were administered to both the control group and the experimental groups one month and one year after the experimental groups had finished their learning phase. Neither of the tests was announced in any way; judging from the participants' comments, the tests came as a complete surprise to all. No seniors volunteered for the experimental groups, so their not taking the one-year test presented no statistical problems.

Statistical Analysis

All tests of significance between two means were done using Student's t -test because it can be used with very small sample sizes. The level of significance chosen is .05, and all values for t have been rounded off to the nearest hundredth. This test of significance, applied to the control group and the experimental group on the pre-test, yielded a $t=1.08$ which is not significant at the .05 level. This means that, statistically, both the control and the experimental groups could be considered as members belonging to one group, although the one hundred eleven members of the control group earned an average of 20.86 on the pre-test, and the fifty-two members of the experimental group only reached an average of 19.4. After the experimental group was divided in two (the 'book' group and the 'computer' group), the two



Graph 2. Comparison of actual data with data generated by estimation curves



resultant groups' scores were also checked. The average for those in the book group was 18.44, and for the computer group it was 20.44; with a t of 0.9, the difference cannot be considered as significant. The average score on the pre-test for all three groups (control, book and computer) was 20.4.

To supplement the observed data, three different types of estimation curves were applied (Linear regression, Logistic curve and Gompertz curve). Estimation curves can be very helpful because, utilizing observed, but often incomplete data, they generate complete data and curves that facilitate comparisons among subjects (see Graphs 1 and 2). In Graph 1 the observed data tells us that the subject reached the 33% level after 12 minutes of study, the 84% level after 161 minutes, the 98% level after 298 minutes, and the 100% level after 447 minutes. We don't have any exact data to tell us how long it took this participant to reach the 50%, 60%, 70%, 80%...levels. Graph 2 shows a different kind of incompleteness; this subject never reached the 100% level. An estimation curve can help by supplying this kind of missing data, although in doing so, we lose some accuracy. Linear regression generates a straight line, but both the logistic curve and the Gompertz curve yield an approximate S-shaped curve which would seem to be more applicable to learning situations; however, in applying these curves to the observed data, no one of them was found to fit in all cases; linear regression provided the best curve fit (matched the observed data best) in 20 cases, the logistic curve in 18 cases and the Gompertz curve in 13 cases. Table 1 provides a summary of the data generated by linear regression, and Table 2 provides the same for the logistic curve (because the Gompertz curve yielded a good fit in less than a third of the cases, it was deleted).

Each Table is divided into five main sections. The first section contains (vertically) the achievement levels written in percents. The second section presents the unaltered estimated mean time for the book group and the computer group to reach each of the achievement levels (30%, 40%, 50%, 60%, 70%, 80%, 90%, 93%, 95% and 98%; the 100% level was not used because estimations at this level tend to be increasingly inaccurate). The third section (a) presents the same type of data, but the scores for the four slowest performing participants (one member of the book group and three from the computer group) have been deleted. The fourth section (A) presents the data corrected for lost time; a correction factor of thirty minutes was used. Those in the computer group first had to learn how to use the computer program; therefore, it took the computer group longer to get started learning the actual pairs. and the fifth section (aA) of each of the Tables incorporates the corrections from both the second and third sections together. The value for t for each of the compared mean times (book and computer) appears to the left of them. Values marked with an '*' indicate that the compared mean times are not significantly different statistically.

Using the data from Tables 1 and 2, Table 1a-2a shows the differences between the estimated mean learning times at each level computed. The format is similar to that of Tables 1 and 2, but as only the differences between the means appear, it is now easy to see that the differences in the amount of time needed to reach each level increase as we approach the 100% level. Even after the scores for the abnormally slow students have been deleted and the learning time for the computer group has been compensated for, the computer group is still 86 minutes behind the book group at the 98% level where the average estimated time difference, using corrected and uncorrected data together for the linear regression, is 123 minutes. Averaging the data and using the logistic curve, it is 271 minutes.

TABLE 1. Linear Regression, estimated mean learning time in minutes

Achievement Level in %	Linear Regression			Linear Regression (a)			Linear Regression (A)			Linear Regression (aA)		
	t value	estimated mean time		t value	estimated mean time		t value	estimated mean time		t value	estimated mean time	
		book	comp		book	comp		book	comp		book	comp
98	3.75	363	523	4.2	349	464	3.05	363	493	3.16	349	435
95	3.83	349	505	4.38	335	448	3.09	349	475	3.22	335	418
93	3.78	339	491	4.3	326	435	3.03	339	461	3.12	326	405
90	3.83	324	471	4.39	311	418	3.05	324	441	3.16	311	388
80	3.96	275	405	4.65	263	360	3.05	275	375	3.21	263	330
70	3.95	230	340	4.69	220	302	2.87	230	310	2.98	220	273
60	3.96	183	274	4.63	174	245	2.66	183	244	2.68	174	215
50	3.63	143	209	4.23	135	188	1.97*	143	179	1.80*	135	158
40	3.17	98	143	3.52	92	130	1.07*	98	113	0.73*	92	100
30	1.05*	67	78	1.3*	63	73	1.3*	67	54	2.12	63	46

a data after deletion of the scores from the four slowest participants

A data corrected for extra time needed by the computer group

aA data after combining the above two

* marks the levels at which there was no significant difference in the means

TABLE 2. Logistic Curve, estimated mean learning time in minutes

Achievement Level in %	Logistic Curve			Logistic Curve (a)			Logistic Curve (A)			Logistic Curve (aA)		
	t value	estimated mean time		t value	estimated mean time		t value	estimated mean time		t value	estimated mean time	
		book	comp		book	comp		book	comp		book	comp
98	4.86	382	706	5.4	365	612	4.41	382	676	4.75	365	582
95	4.86	319	584	5.47	305	508	4.31	319	554	4.66	305	478
93	4.86	296	538	5.49	282	468	4.26	296	508	4.61	282	439
90	4.85	270	488	5.53	257	426	4.19	270	458	4.55	257	396
80	4.81	217	384	5.59	206	336	3.95	217	354	4.32	206	308
70	4.75	181	315	5.58	172	278	3.68	181	285	4.01	172	248
60	4.61	152	258	5.5	143	229	3.28	152	228	3.6	143	200
50	4.46	125	206	5.45	118	188	2.74	125	176	3.23	118	160
40	3.76	98	154	4.41	92	138	1.73*	98	124	1.55*	92	108
30	1.87*	76	97	1.87*	70	86	0.74*	76	67	1.59*	70	56

a data after the deletion of the scores from the four slowest participants

A data corrected for extra time needed by the computer group

aA data after combining the above two

* marks the levels at which there was no significant difference in the means

TABLE 1a-2a. Differences between estimated mean times

Achievement level in %	Linear Regression	Linear Regress. (a)	Linear Regress. (A)	Linear Regress. (aA)	Logistic Curve	Logistic Curve (a)	Logistic Curve (A)	Logistic Curve (aA)
98	160	115	130	86	324	247	294	217
95	156	113	126	83	265	203	235	173
93	152	109	122	79	242	186	212	157
90	147	107	117	77	218	169	188	139
80	130	97	100	67	167	130	137	102
70	110	82	80	53	134	106	104	76
60	91	71	61	41	106	86	76	57
50	66	53	36*	23*	81	70	51	42
40	45	38	15*	8*	56	46*	26*	16*
30	11*	10*	-7*	-17*	21*	16*	-10*	-14*

a data after the deletion of the scores from the four slowest participants

A data corrected for extra time needed by the computer group

aA data after combining the above two

* marks the levels at which there was no significant difference between the means

— indicates where the mean estimated time is lower for the computer group

Consulting Table 3, which contains the observed mean learning times, we can see that in all but two cases, the observed mean times are lower than those of either estimation type in Tables 1 and 2. In addition, the observed differences between the book group and the computer group at the 94% to 100% level (139 and 109 (A)) are also lower than averaging the appropriate values appearing in Table 1a-2a (158 and 128 (A) for linear regression, and 295 and 265 (A) for the logistic curve). However, the differences in observed mean times at the

TABLE 3. Observed and interpolated mean learning time in minutes

Achievement Level in %	t value	Observed ^o mean time		t value	Observed ^o (A) mean time	
		book	comp		book	comp
94-100	3.31	330	469	2.6	330	439
80	4.27	253	413	3.47	253	383
50	3.32	135	201	1.83*	135	171

A data corrected for extra time needed by the computer group

^o includes some interpolated times

* marks the level at which there was no significant difference in the means

80% level are higher than the values estimated by linear regression. At the 50% level, both the differences in the observed mean times are equal to those estimated by linear regression, but lower than those estimated by the logistic curve. The inescapable conclusion: at least above the 50% level, there is a significant difference between the two groups; the book group learned faster than the computer group.

One month after the experimental groups had completed their learning phase, all participants (including the control group, which had lost eighteen members) were given another pairing test. As expected, the experimental group (both book and computer group together) did much better on the test. The mean for the control group was 22.9% and the mean for the experimental groups was 74.6%; with a t of 28.6, the difference in their means is significant at any level chosen. If we compare the control's mean of 22.9% with the experimental groups' mean for the last tests of their learning phase, 87.1%, there is, of course, a greater difference: 64.2 percentage points as opposed to 51.7. Looking at the differences between their scores on the pre-test and post-test (mentioned just above), we can see that the experimental groups gained an average of 55.4 points, whereas the control gained an average of only 2.8.

Turning our attention to the post-test performance of the experimental groups, let us note some interesting differences. The mean score for the book group was 77.6%, and for the computer group it was 75.5%, a difference of 2.1, and a t of 0.58, which is not significant at any level. At the end of their learning phase, the means were 94.9% and 86.7% for the book and computer groups, respectively. Comparing this with their respective mean scores on the post-test leads us to the conclusion that the computer group was able to retain significantly more than the book group. For the book group, there was a mean loss of 17.3 percentage points, while for the computer group, there was only a mean loss of 11.2. With a t of 2.21, this is significant at the .05 level.

In addition to the post-test administered one month after the completion of the learning phase, another post-test was given to all the participants (except for the graduated seniors) one year after the learning phase was finished. With respective means of 15.6% and 37.9%, the control group was still behind the experimental groups. This difference is significant at any level, but certainly comes as no surprise. What is interesting is the fact that after one year

of no study, the computer group slightly outperformed the book group: 38.7% to 37.5% which, on the surface, is not statistically significant. However, computing the mean loss in retention, we can see that the book group suffered a 57.3 loss compared to 45.7 for the computer group. This difference of the means (11.6 percentage points with a $t=4.1$) is statistically significant at any level.

The hundred-plus pages of computer-generated data, the computer programs, the students' notebooks, and the tests are all available for inspection at the author's university office.

Conclusion

The safest conclusion that can be drawn from the research presented herein is that much more research needs to be done not only to determine the differences between older forms of learning and computer-assisted learning, but also to determine in what areas computer-assisted learning can be used effectively. If the data from this research is not misleading, then we can also conclude that, while using a book for rote learning tasks may be faster than using a computer, utilizing a computer may result in longer retention of the material studied. This conclusion, however, needs to be qualified because those in the computer group spent more time in the learning phase, that is, it took the computer group longer to reach their mean of 86.74% than it did for the book group to reach its mean of 94.85%. Spending more time in the learning phase may naturally result in greater retention. In addition to this, the computer program was set up to present the items in a quasi-random fashion; no one could choose which items he wanted to study, thus each participant in the computer group was subjected to a higher level of repetition than those in the book group who just studied what they thought they still didn't know. Some of the students complained of this feature of the computer program (they wanted to study only the words and idioms that they didn't know), and some thought that they had forgotten what they had learned because they couldn't immediately recall an item after the program went on to the next question (because of the mechanics of the program, they couldn't go back and check). And in addition, a number of members of the computer group pointed out that the noise level in the computer room was greater than in the study room for the book group (this interference was not compensated for in the statistical analysis). No doubt, the members of the computer group will be surprised to know that they did significantly better in retention than the book group.

Acknowledgments

I would like to express my sincere thanks to all the English students at Osaka University of Foreign Studies who gave of their energy and free time to work on this research. Needless to say, without their participation, this research couldn't have been completed. In addition, I would like to express my special thanks to Miss Noriko Aoi and Miss Aki Naito who greatly helped with the organization of the original raw data; their efforts made it much easier for me

to produce this report. And again I would like to thank Mr. Masaki Manabe and Mr. Mitsuo Yoshida for their help in taming my computer.

APPENDIX A

List of Pairs Used in Learning Program

1. admit	deny
2. a fast buck	hard earned money
3. afraid of nothing	afraid of one's own shadow
4. agile, graceful	awkward, clumsy
5. ahead of the game	behind the eight-ball
6. all ears	in a zone
7. an egghead	a turkey
8. appease	irritate
9. apple of one's eye	thorn in one's side
10. a rip-off	a steal
11. back down from	stand up to
12. bad egg	good egg
13. barbarous	civilized
14. bear a grudge	carry a torch
15. beat around the bush	come to the point
16. be at loggerheads	see eye to eye
17. be dependent on	be independent of
18. be out of it	be with it
19. beyond one's reach	within one's reach
20. big time	penny ante, small time
21. bought for an arm and a leg	bought for a song
22. bury the hatchet	cross swords
23. butter someone up	chew someone out
24. call it quits, give up, throw in the towel	carry on, hang in there, keep pushing
25. catch sight of	lose sight of
26. change one's tune	stick to one's story
27. chicken-livered	lion-hearted
28. clam up	rattle on
29. close down, shut down	open up
30. come to	pass out
31. complicated, complex	simple
32. count me in	count me out
33. courageously	cowardly

34. crazy as a loon	sober as a judge
35. criticize	praise
36. crushing failure, whopping failure	howling success, roaring success
37. deposit	withdraw
38. die out	live on
39. drunk	sober
40. dry behind the ears	wet behind the ears
41. dumb as an ox	smart as a fox
42. ear to the ground	head in the sand
43. easy way	hard way
44. eat like a bird	eat like a horse
45. end up	start out
46. even number	odd number
47. exotic	indigenous
48. extroverted, outspoken	introverted, withdrawn
49. fair play	foul play
50. fall apart, go to pieces	pull oneself together, put together
51. fall short	surpass
52. feet on the ground	head in the clouds
53. first class	fourth class, last class
54. get it all together	lose it
55. get the brushoff, get the cold shoulder	get the eye
56. good loser	sore loser
57. hale and hearty	sickly
58. hard sell	soft sell
59. have a bad time	have a good time
60. heart of gold	heart of stone
61. hit the deck	hit the hay
62. hold one's temper	lose one's temper
63. holy terror, little beast, little devil	little angel, little darling
64. hurry up, make it snappy	slow down, take it easy
65. in character	out of character
66. in one's element	out of one's element
67. in public, openly, out in the open	on the Q.T., on the sly, secretly
68. in shape	out of shape
69. inside out	right side out
70. in the black	in the red
71. in the dark	in the know, in the light
72. keep one's head	lose one's head
73. keep track	lose track

74. know-it-all	know-nothing, smart aleck
75. live beyond one's means	live within one's means
76. live from hand to mouth	live high off the hog
77. long and drawn out	short and sweet
78. long haul	short haul
79. look down on	look up to
80. lose heart	take heart
81. make good sense, make sense	make no sense
82. make one's mouth water	turn one's stomach
83. master	servant, slave
84. neat	sloppy
85. off by a mile, way off	on the nose, on target, right on
86. off one's guard	on one's guard
87. on one's bad side	on one's good side
88. perk up	peter out
89. play down	play up
90. quick as greased lightning, quick as lightning	slow as molasses in January
91. right side of the tracks	wrong side of the tracks
92. right side up	upside down
93. soft touch	tight wad
94. stay on the fence	take sides
95. take it lightly	take it seriously
96. take the bull by the horns	turn tail
97. the bottom of the barrel	the cream of the crop
98. to be up the creek	to have smooth sailing
99. tune in	tune out
100. walk in	walk out